# Semantic Ambiguity in Expert Systems: The Case of Deterministic Systems

Daniel E. O'Leary

School of Business
University of Southern California
Los Angeles, CA 90089-1421
213-740-4856
oleary@mizar.usc.edu

**Abstract.** Typically with expert systems, the user is asked to assess or categorize evidence (E), based on data (D). The system is then responsible for matching that evidence (E) to the appropriate hypothesis (H). Throughout, the user provides the characterization of data as evidence for the system.

Unfortunately, different users may develop different categorizations of data. Based on different points of view, different users may categorize D as E or E'. This is referred to as semantic ambiguity. Such potential differences can have a major impact on the results recommended by expert systems.

This paper presents empirical results that indicate that there can be substantial ambiguity about the categorization of D as E or E'. In one situation, 50% of the subjects categorized data as E and 50% as E'. This paper also develops a probability model that is used to investigate semantic ambiguity in deterministic systems (no weights on the rules).

## 1 Introduction

Typically with expert systems, the user is asked to assess or categorize evidence (E), based on a set of data (D). The system is then responsible for matching that evidence (E) to the appropriate hypothesis (H). Throughout, the user interprets the data as evidence for the system.

For example, in the following rule [1, p. 512]), users of the expert system are asked to determine if the "... market followed an upward trend recently...."

If 1. the client's income tax bracket is 50%
2. the market has followed an upward trend
Then, there is evidence that the area of the investment should be high technology.

In this case, actual market performance is D, "upward trend" is E and "the area of the investment being high technology" is H.

Unfortunately, developing categorizations of data is not a straightforward task. Few probably would argue with the evidence categorization of the data (2850, 2900 and 3000) as following an "upward trend." However, the matching between the data



and the meaning of terms of concern, may not be so clear in all situations. For example, does the data set (2900, 3000, 2875), follow an upward trend?

The ambiguity in the matching between E and D is referred to here as "semantic ambiguity." For each E and D there is a distribution associated with the interpretation of D as E and D as E' (not E).

### 1.1 Contributions of this Paper

This paper presents empirical work that suggests, depending on D and E, different users are likely to provide very different evidence assessments to the same situations. In some cases, given D, users provide the system with the assessment E, while other users provide the assessment E'.

This paper also develops an analytic model that allows us to investigate semantic interpretation. In the case where there are no weights on the rules, we can characterize and capture the probability that different assessments will be made by different users or developers for the same data set. For example, if the probability of H, given E is greater than the probability of H', given E then we could argue that only the rule linking E to H should be built in the deterministic system. However, for each user of the system to derive the same system response, that decision rule requires that under similar circumstances, each user should evaluate D as E. Thus, semantic uncertainty could result in different users, under the same circumstances, generating different system responses. As a result, if the gap between the probability of H, given E, and the probability of H, given E and D is large or costly, then it may be preferable to have the system categorize the data as evidence, rather than depending on the user.

## 2 Semantic Ambiguity

This paper is concerned with situations that can be characterized as follows (but also can be generalized). Assume that the user is confronted with data, D, or D' (not D) and is required to categorize that data as either E or E'. The expert system takes the user's assessment, E or E', as input, and relates that evidence deterministically with some hypothesis H. As a result, we are modeling rules of the type "If E then H," given the user assesses D as E.

P(x) is used to denote the probability of x, while $P(x|y)$ is the conditional probability of x given y. Throughout either D or D' occur and neither can be further partitioned. Either E or E' are the resulting categories, and neither can be further partitioned.

It will be said that there is semantic ambiguity in a rule "If E then H" if the probability of E given D is less than one or greater than 0 ($1 > P(E|D) > 0$). Thus, given D and no semantic ambiguity, we know that the rule "If E then H," either will or will not be executed for all users.

The problem of semantic interpretation is where one user may view D as E, with complete certainty, while another user may view D as E', also with complete certainty. (Or a user at one point in time categorizes D as E, with certainty, and at another point in time, categorizes D' as E, also with complete certainty.) At any rate, once the interpretation is made by the user, the system takes that interpretation as the only evidence, even though different users may provide completely different assessments E for the same D.

## 3 An Empirical Analysis

This section presents evidence that there can be substantial semantic ambiguity present in an expert system. Particular concern is given to the case where only two choices are allowed, e.g., "yes" or "no." Such binary choice situations are common in deterministic expert systems and other analysis tools, such as decision trees. Binary choices force the user to make crisp decisions. Unfortunately, these crisp decisions are often in situations where there is substantial ambiguity. The extent of that ambiguity is reflected in large percentages of the users choosing each alternative.

### 3.1 Tax Advisor

Tax Advisor [4] was investigated for potential semantic ambiguity. The following rules have a number of words that present the potential for semantic ambiguity. These rules were used to generate a test instrument that was used to perform an empirical analysis of the impact of semantic ambiguity in expert system situations.

**Rule 142**

If: 1) The *degree of financial management, experience,* and *ability of the client and/or spouse is poor,*
2) The client and/or spouse have assets that he/she is *willing to dispose of*
Then: It is definite that the client and/or spouse should invest in *low management investments.*

**Rule 208**

If: 1) The client and/or spouse is *willing* to consider a *low risk,* fixed dollar investment with a fixed return
2) Client and/or spouse does not have the time and/or *financial skills* to select his/her own security issues ....

**Rule 175**

If: ... the client has an estate that is *not liquid enough* to pay death taxes, administration, medical and funeral expenses and the immediate living expenses of the family without the sale of assets ....

Some of the ambiguous judgments are underlined in the above rules. In many cases the rules are quite extensive. As a result, these rules have been edited for presentation.



### 3.2 Questions and Subjects

The above rules force the users into a number of different judgments. For example, in two of those rules the user would be asked to assess "enough." What is enough insurance or what is enough liquidity? These answers are likely to vary between users and situations.

Subjects were faced with binary choices, rather than a graded choice: e.g., are the financial management skills "poor" or "not poor." An instrument was developed that required the subjects to present answers that are typical of bianry choice expert systems.

Questions were taken from the rules in Tax Advisor, discussed above, to investigate ambiguity in tax and accounting systems. The questions used in the empirical study are summarized in Table 1.

**Table 1.** Tax Planning Questions

A tax advisor must determine certain facts about the client in order to make recommendations related to tax planning. The following questions will present certain facts and you will be expected to rate the client (investment) attributes based on your own judgments.

**Question 1 (Financial Skills)**

A. The client is a well-respected attorney who specializes in trusts and estates.

His degree of financial management experience is

Poor Good

This client has the financial skills to select his own security issues.

No Yes

B. The client is a recently widowed housewife. She has been married for 28 years and has never held a compensated job. She has handled the family funds, including liquid investments totalling $250,000. She was president of the local grammar school PTA for five years.

Her degree of financial management experience is

Good Poor

This client has the financial skills to select her own security issues

No Yes

C. The client is a vice president of a large engineering firm in which he is a limited partner. He holds an M.S. in engineering as well as an M.B.A., both of which he earned when he was a young man.

His degree of financial management experience is

Good Poor

This client has the financial skills to select his own security issues

No Yes

**Question 2 (Asset Questions)**

A. The client currently holds assets consisting of mutual funds ($30,000), coupon bonds ($30,000), a piece of developed residential property (net fair market value of $800,000) and a family home of (fair market value of $700,000, encumbrance of $600,000).

These assets require a substantial degree of management

No Yes

The client has an estate that is liquid enough to pay death taxes, administration, medical and funeral expenses and the immediate living expenses of his spouse and two children

No Yes

B. The client currently holds assets consisting of a grantor trust which holds several variable annuities. These assets require a substantial degree of management

No Yes

The client has an estate that is liquid enough to pay death taxes, administration, medical and funeral expenses and the immediate living expenses of his spouse and two children

No Yes

C. The client currently holds assets consisting of a Southern California coin operated laundry mat and neighboring undeveloped land. These assets require a substantial degree of management

No Yes



The client has an estate that is liquid enough to pay death taxes, administration, medical and funeral expenses and the immediate living expenses of his spouse and two children

No Yes

D. The client currently has an estate which consists of several five year second trust deeds secured by residential real estate. These deeds have been issued throughout the past five year period. These assets require a substantial degree of management

No Yes

The client has an estate that is liquid enough to pay death taxes, administration, medical and funeral expenses and the immediate living expenses of his spouse and two children

No Yes

---

The instrument was developed and pilot tested on a faculty member and a Ph. D. student. Then the instrument was directly administered to 18 students in an advanced class in the Master of Tax program at the University of Southern California. Students are appropriate subjects since in many situations expert systems are used to delegate expertise down to lower levels in the organization. Virtually all these students were employed as professionals at the conclusion of the term in which this study was made.

### 3.3 Findings

The results are summarized in Table 2. An analysis of the data indicates that there apparently was substantial ambiguity. The greatest ambiguity was in question 1-C, where the respondents were equally divided between the two possible responses. A number of other situations also provided substantial semantic ambiguity. In addition, in no situation presented were the subjects unanimous in their assessment of the situation. The minimum split was with question 2 D (22.2% and 77.8%).

**Table 2.** Summary of Results on Tax Questions

| Concept/Question | Response % | | | |
|---|---|---|---|---|
| | Poor | Good | No | Yes |
| **Question 1 - "Financial Skills"** | | | | |
| A. "Financial Management" | 33.3 | 66.7 | | |
| A. "Select Securities" | | | 38.3 | 61.7 |
| B. "Financial Management" | 27.9 | 72.1 | | |
| B. "Select Securities" | | | 55.6 | 44.4 |
| C. "Financial Management" | 55.6 | 44.4 | | |
| C. "Select Securities" | | | 50.0 | 50.0 |

**Question 2 - "Asset Management"**

| | | | | |
|---|---|---|---|---|
| A. "Degree of Management" | 66.7 | 33.3 | | |
| A. "Liquid Enough" | | | 55.6 | 44.4 |
| B. "Degree of Management" | 50.0 | 50.0 | | |
| B. "Liquid Enough" | | | 61.7 | 38.3 |
| C. "Degree of Management" | 55.6 | 44.4 | | |
| C. "Liquid Enough" | | | 72.1 | 27.9 |
| D. "Degree of Management" | 22.2 | 77.8 | | |
| D. "Liquid Enough" | | | 66.7 | 33.3 |

Such differences would guide the users to different parts of the rule base. Different users of the expert system would be provided with different answers to the same questions.

# 4 Single and Multiple Interpretations

In the development of an expert system there are at least two perspectives that guide the process of when a rule should be added to the system. Those perspectives derive from the different interpretations of the relationships between, D, E and H. First, it might be argued that there is a single correct interpretation of D as E, otherwise, expertise is lost. An expert would categorize the data in one way and expect the system to process it accordingly. In that situation, in order to capture expertise, the system must be based on a single correct interpretation of the data.

However, the empirical study in the previous section indicates that users of expert systems are likely to make multiple evidence categorizations of the same data. From a design and development perspective, there would be no "correct" answer. Because of the resulting ambiguity, generally systems are *not built* assuming multiple interpretations of the data. However, when they are *used*, there may be multiple interpretations of the same data.

Thus there is interest in determining when to add a rule to an expert system, given that it is assumed that there will be either a single interpretation (no semantic ambiguity) or multiple interpretations (semantic ambiguity).

## 4.1 Single Interpretation: System Design

If the expert thinks the data has an "upward trend," then it is likely that it is assumed that all must interpret the data in the same manner. Assuming a single interpretation is consistent with the notion of expertise: the expert views the world in a given manner and the system should be built to reflect that view.

If we assume a single interpretation in model design, then ideally a rule "If E then H," should be added to the system if the evidence *and* the data indicate that it is the appropriate rule. One approach to the choice of rules is to employ a probability-based decision rule. In particular, call the "evidence and data rule" the following: if $P(H|E,D) \geq P(H|x,y)$, for $(x,y)=(E,D')$, $(E',D)$ and $(E',D')$ then add the rule "If E then H". Although developers may not use this decision rule explicitly



to design an expert system, the single interpretation assumption, results in the *effective* assumption of that decision rule.

## 4.2 Multiple Interpretations: System Use

Alternatively, based on the above empirical results, that users of expert systems may categorize evidence in multiple ways. There may be two feasible interpretations of D *and* D' as E (and, thus, D and D' as E'). In fact much of human communications seems to focus on these alternative interpretations.

The use of an expert system differs from the design in the relationship of the data and the evidence. In the case of semantic ambiguity, different users may employ either D or D' to develop E, which leads to H. Thus, the use of an expert system indicates that a rule "If E then H" will be employed if $P(H|E,D,D') \geq P(H|E',D,D')$.

## 4.3 Relationship Between Use and Design

Thus, it is assumed that the designer/developer assumes a single correct interpretation in the design and development of the system. Similarly, it is assumed that users may have multiple interpretations in the use of the system. As a result, we can see that development decision rules and use decision rules are not the same. In particular, $P(H|E,D)$ does not equal $P(H|E,D,D')$. Thus, probability models indicate that use and design perspectives result in different systems.

As a result, as the distance of $P(H|E,D)$ and $P(H|E,D,D')$ increases, we are likely to find that the use of expert systems that require the user to provide data will become more ineffective. As that distance increases (or if the cost is large), it is likely that other approaches to evidence categorization would be more appropriate. For example, rather than having the user provide category estimates from data, the system might be used to provide those estimates. In the case of determining "increasing," a computer estimate of the slope might be used.

# 5 Uniqueness Properties of Design Decision Rules

An important property of designing expert systems is that the decision rules for adding rules results in unique systems. This section provides uniqueness proofs for two such decision rules.

## 5.1 The Evidence Rule

The evidence rule can be characterized as follows: If $P(H|E) \geq P(H|E')$ then add the rule "If E then H". The use of the evidence rule might occur in those situations where it is felt there may be no ambiguity of interpretation or where there is no typical data that would be associated with the evidence. The evidence rule may be used in those situations where there are a set of heuristic rules or departmental rules, since those rules are generally coded without reference to specific occurrences (data).

The evidence rule will result in a unique set of rules (excluding ties at equality). There will be at most one rule with either E or E' and H. Similarly, there will be at most one rule with either E or E' and H'. Further, each of E, E',H and H' will be in at most one rule, with respect to those variables. This is guaranteed by the following theorem.

*Theorem 1* -- Uniqueness

$P(H|E) \geq P(H|E')$ iff $P(H'|E) \leq P(H'|E')$.

*Proof*

$P(H|E) \geq P(H|E')$, implies that $1 - P(H|E) \leq 1 - P(H|E')$. But that implies that $P(H'|E) \leq P(H'|E')$. The other direction has a similar proof.

### 5.2 The Data Rule

The data rule can be characterized as follows: If $P(H|D) \geq P(H|D')$ then add the rule "If E then H". This approach might be used in those situations where the knowledge acquisition approach employs data to capture the concepts associated with the evidence. In this case the evidence categories may not be understood, but the data may have clear implications, e.g., in the situation of case-based reasoning. For example, bankrupt firms may have similar data characteristics that are arbitrarily labeled as a type of evidence, such as "liquidity" or "available cash."

The data rule also will result in a unique set of rules (excluding ties at equality). There will be at most one rule with either D or D' and H. Similarly, there will be at most one rule with either D or D' and H'. This is guaranteed by the following theorem.

*Theorem 2* -- Uniqueness

$P(H|D) \geq P(H|D')$ iff $P(H'|D) \leq P(H'|D')$.

## 6 Summary and Extensions

This section provides a brief summary of the paper and a discussion about some of the possible extensions of the research presented in this paper.

This paper developed a model of semantic interpretation and studied the impact of that interpretation on deterministic models. Both empirical and analytical approaches were used to study the problem. The empirical portion of the paper presented evidence for the existence of semantic ambiguity. In one case 50% of subjects categorized data as E, while 50% categorized that same data as E'.

The analytical analysis pursued the differences between systems due to a single interpretation assumption and a multiple interpretation assumption. The use and development of expert systems each seem to require a different one of those



assumptions. As the gap between those assumptions increases it appears important to shift the evidence categorization process from the user to other sources.

The approach used in this paper can be extended in a number of ways. First, only a single piece of evidence and a single hypothesis were considered. The results presented here can be extended to the integration of multiple pieces of evidence.

Second, the results can be extended to probabilistic models. Bayesian systems, such as AL/X [2,3] can be extended to account for ambiguity. In addition, the results can be extended to an the use in influence diagrams.

Third, the results presented in this paper indicate that it is critical for the development of approaches other than rules, in the design of expert systems. In particular, it is apparent that dialogues between the system and the user are probably necessary to mitigate the impact of semantic ambiguity.

## References

1. Davis, R. and Buchanan, B., "Meta Level Knowledge," in Buchanan, B. and Shortliffe, E., Rule-based Expert Systems, Addison-Wesley, Reading, Massachusetts, 1985.

2. Duda, R., Gaschnig, J., and Hart, P., "Model Design in the Prospector Consultant System for Mineral Exploration," in D. Mitchie, (Ed.) Expert Systems for the Microelectronic Age, Edinburgh, Edinburgh University Press, 1979.

3. Duda, R., Hart, P. and Nillsson, N., "Subjective Bayesian Methods for Rule-based Inference Systems,"National Computer Conference, pp. 1075-1082, 1976.

4. Michaelsen, R., "A Knowledge-based System for Individual Income and Transfer Tax Planning," Unpublished Ph. D. Dissertation, University of Illinois, 1981.